AD-A035 929 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER --ETC F/G 12/1
COMPUTATIONAL COMPLEXITY OF ONE-STEP METHODS FOR SYSTEMS OF DIF--ETC(U)
SEP 76 A G WERSCHULZ N00014-76-C-0370
UNCLASSIFIED NL

| OF |
AD
A035929

END
DATE
FILMED
3-77

Computational Complexity of One-Step Methods for Systems of Differential Equations

Arthur G. Werschulz

September, 1976

Department of Mathematics
Carnegie-Mellon University
Pittsburgh, PA 15213

Abstract: The problem is to calculate an approximate solution of an initial value problem for an autonomous system of N ordinary differential equations. Using fast power series techniques, we exhibit an algorithm for the p^{th}-order Taylor series method requiring only $O(p^N \ln p)$ arithmetic operations per step as $p \to +\infty$. (Moreover, we show that any such algorithm requires at least $O(p^N)$ operations per step.) We compute the order which minimizes the complexity bounds for Taylor series and linear Runge-Kutta methods, and show that in all cases, this optimal order increases as the error criterion ε decreases, tending to infinity as ε tends to zero. Finally, we show that if certain derivatives are easy to evaluate, then Taylor series methods are asymptotically better than linear Runge-Kutta methods for problems of small dimension N.

This research was supported in part by the National Science Foundation under Grant MCS75-222-55 and the Office of Naval Research under Contract N00014-76-C-0370, NR 044-422.

1. Introduction

Let $\mathcal{D}$ be a set of points in the real N-dimensional linear space $\mathbb{R}^N$, and let $\mathcal{V}$ be a set of operators on $\mathbb{R}^N$, such that the initial value problem of finding a function $x : [0, 1] \to \mathbb{R}^N$ satisfying

$$\dot{x}(t) = v(x(t)) \quad \text{if } 0 < t < 1$$
$$x(0) = x_0 \tag{1.1}$$

has a unique solution for every $(x_0, v) \in \mathcal{D}\times\mathcal{V}$; we assume that x is analytic on [0, 1]. The autonomous form of this system is no restriction, since any non-autonomous system may be made autonomous by increasing the dimension of the system by one.

In Werschulz [76], we looked at the computational complexity of using one-step methods to generate an approximate solution to (1.1) on an equidistant grid in the sense of Stetter [73]; that is, the methods considered computed approximations x_i to x(ih) by the recursion

$$x_{i+1} = x_i + h\,\varphi(x_i, h) \qquad (0 \le i \le n-1,\ n = h^{-1}), \tag{1.2}$$

where $h = n^{-1}$ is the step-size of a grid with n points, and φ is the increment function (Henrici [62]) for the method. (To be brief, we will refer to "the method φ.") In that paper, we discussed the problem of optimal order and minimal complexity for rather general classes of one-step methods.

In this paper, we will use the techniques and results of Werschulz [76] to analyze the complexity of using Taylor series methods and linear Runge-Kutta methods to generate approximate solutions whose error does not exceed ε. The model of computation, error measure, and complexity measure to be used are described in Section 2, as well as the relevant results from Werschulz [76].

We discuss the complexity of Taylor series methods in Section 3. Using the fast power series techniques of Brent and Kung [76], we show that $O(p^N \ln p)$ arithmetic operations suffice to compute the p^{th}-order Taylor series approximation; moreover, we show that $O(p^N)$ operations are necessary. In Section 4, we discuss the complexity of linear Runge-Kutta methods. In both Sections, we compute lower and upper bounds on the complexity using a fixed method of given order; these results are then used to compute optimal orders which minimize these complexity bounds. We show that in all cases, the optimal order increases as ϵ decreases, tending to infinity as ϵ tends to zero.

Finally, we compare these two classes of methods in Section 5, where we show that if the partial derivatives of v are easy to evaluate, then Taylor series methods are asymptotically better (as ϵ tends to zero) than linear Runge-Kutta methods for problems of small dimension N.

2. Preliminary Results

Before proceeding any further, we will establish some notational conventions. Let $\mathfrak{X}$ be an ordered ring; then $\mathfrak{X}^+$ and $\mathfrak{X}^{++}$ respectively denote the nonnegative and positive elements of $\mathfrak{X}$. (This is used in the cases $\mathfrak{X} = \mathbb{R}$, the real numbers, and $\mathfrak{X} = \mathbb{Z}$, the integers.) The symbol ":=" means "is defined to be," while "≡" means "is identically equal to." We use "I" to denote the unit interval [0, 1]. The symbol "∇" is used to denote the gradient of a mapping. The notations "$x \downarrow a$" and "$x \uparrow a$" are used to indicate one-sided limits, as in Buck [65]. Finally, we write "$(a.b)_c$" to indicate the c^{th} part of equation (a.b), as in Gurtin [75].

We next describe the model of computation to be used. We assume only that all arithmetic operations are performed exactly in $\mathbb{R}$ (i.e., infinite-precision arithmetic) and that for any algorithm to be considered for the solution of (1.1), a set of procedures is given for the computation of any information about v required by that algorithm. (For instance, with Runge-Kutta methods, we must be able to compute v at any point in its domain.)

In addition, we must pick an error measure, so that we may measure the discrepancy between the approximate solution produced by φ (via (1.2)) and the true solution. For the sake of definiteness, we use the global error

$$\text{(2.1)} \qquad e_G(\varphi,h) := \max_{0 \le i \le n} \|x(ih) - x_i\| ,$$

where $\|\cdot\|$ is a norm on $\mathbb{R}^N$. Other error measures may be used, such as the local error per step e_L and the local error per unit step e_{LU} (see Henrici [62] and Stetter [73] for definitions); this would involve only a slight modification of the results contained in the sequel.

We finally describe the complexity measure to be used. Let $\Phi = \{\varphi_p : p \in \mathbb{Z}^{++}\}$ be a basic sequence in the sense of Werschulz [76]; that is, there exist functions $\kappa : \mathbb{R}^+ \times I \to \mathbb{R}^+$ and $\kappa_L , \kappa_U : \mathbb{R}^+ \to \mathbb{R}^+$ such that

$$\sigma_G(\varphi_p,h) = \kappa(p,h)\, h^p \quad \text{for } h \in I \text{ and } p \in \mathbb{Z}^{++}, \tag{2.2}$$

where

$$0 < \kappa_L(p) \le \kappa(p,h) \le \kappa_U(p) < +\infty \quad \text{for } h \in I . \tag{2.3}$$

We say that φ_p has order p. This is a slight extension of the definition of order given in Cooper and Verner [72]; the function κ_L introduced here is necessary and sufficient for the "order" of a method to be unique. (For the sake of exposition, we assume that κ_L and κ_U are analytic on $\mathbb{R}^+$, and that $\lim_{p\to 0} \kappa_L(p)^{1/p}$ and $\lim_{p\to 0} \kappa_U(p)^{1/p}$ exist and are positive real numbers; this will always be the case in the examples we consider.) Then we will be interested in the total number of arithmetic operations $C(p,\alpha)$ required to guarantee that

$$\sigma_G(\varphi_p,h) \le \varepsilon := e^{-\alpha} , \tag{2.4}$$

for a given p and a given α. (Here e is the base of the natural logarithms.) We suppose that $0 < \varepsilon < 1$, so that α is positive. Clearly, α increases as ε decreases, and α tends to infinity as ε tends to zero.

In the methods we consider, we may write

$$C(p,\alpha) = n\, c(p) , \tag{2.5}$$

where n is the minimal number of steps required and the cost per step c(p) is the number of arithmetic operations required for the method of order p. As in Traub and Woźniakowski [76], we shall express the cost per step associated with φ_p in the form

$$c(p) := e(\mathfrak{N}_p(v)) + d(p) . \tag{2.6}$$

Here $\mathfrak{N}_p(v)$ is the information about v required to perform one step of φ_p, and we

write $e(\mathfrak{N}_p(v))$ for the informational cost of φ_p; we call d(p) the combinatory cost of φ_p.

Note that we explicitly indicate the dependence of $\mathfrak{N}_p$ on v, so that we may compare the cost of (say) an evaluation of v with a scalar arithmetic operation. Basically, $e(\mathfrak{N}_p(v))$ measures the cost of getting new data about v required by φ_p, while d(p) measures the cost of combining this new data to get an approximate value of the solution at a new point. For example, Euler's method in $\mathbb{R}^N$

$$x_{i+1} = x_i + h\, v(x_i)$$

has informational cost $\Sigma_{i=1}^{N} e(v_i)$, where $v_1, \ldots, v_N$ are the components of v, and for any function ω: $\mathbb{R}^N \to \mathbb{R}$, we define

(2.7) $\qquad e(\omega) :=$ cost of evaluating ω at one point .

The combinatory cost is 2N arithmetic operations, i.e., one scalar multiplication and one scalar addition for each of the N components.

We must now face a problem that occurs in almost all areas of complexity theory. The number of operations c(p) required for one step of a p^{th}-order method is usually unknown per se; we only have bounds of the form

(2.8) $\qquad c_L(p) \le c(p) \le c_U(p)$.

That is, $c_L(p)$ is a lower bound on the number of operations required per step, usually derived via theoretical considerations, and $c_U(p)$ is an upper bound on the number of operations required per step, which is derived by exhibiting an algorithm for computing the p^{th}-order method. (In what follows, we shall assume that the functions $c_L, c_U : \mathbb{R}^+ \to \mathbb{R}^+$ are analytic, although this requirement may be greatly weakened. However, this assumption holds for all examples that we consider.)

From the discussion in Section 3 of Werschulz [76], we find that the step-size h must satisfy

(2.9) $$h_U(p,\alpha) \le h \le h_L(p,\alpha),$$

where

(2.10) $$h_L(p,\alpha) := \kappa_L(p)^{-1/p}\, e^{-\alpha/p} \quad \text{and} \quad h_U(p,\alpha) := \kappa_U(p)^{-1/p}\, e^{-\alpha/p}.$$

Using (2.5), (2.8), (2.9), and (2.10), we may find bounds on the complexity $C(p,\alpha)$.

Theorem 2.1: Define

$$C_L(p,\alpha) := f_L(p)\, e^{\alpha/p}, \quad \text{where} \quad f_L(p) := \kappa_L(p)^{1/p}\, c_L(p),$$

and

$$C_U(p,\alpha) := f_U(p)\, e^{\alpha/p}, \quad \text{where} \quad f_U(p) := \kappa_U(p)^{1/p}\, c_U(p).$$

Then

(2.11) $$C_L(p,\alpha) \le C(p,\alpha) \le C_U(p,\alpha).$$

Proof: See Theorem 3.1 of Werschulz [76]. ∎

Thus we have bounds on the complexity of using φ_p to compute an approximate solution satisfying (2.4). We now wish to consider the problem of optimality. Define

(2.12) $$C^*(\alpha) := \inf \{C(p,\alpha): \varphi_p \in \Phi\}.$$

We are interested in bounds for $C^*(\alpha)$ under reasonable assumptions about f_L and f_U. We first suppose that

(2.13) $$f_L(p) > 0 \quad \text{and} \quad f_U(p) > 0 \quad \text{if} \quad p > 0$$

and

(2.14) $$\lim_{p\uparrow\infty} f_L(p) = \lim_{p\uparrow\infty} = +\infty.$$

Assumption (2.13) is that there is no method whose cost per step is zero, while (2.14) essentially means that the "better" a method is (*i.e.*, the higher its order is), the more we should expect to pay for its use.

Using the techniques of elementary calculus, we find that a necessary condition for p to minimize $C_L(\cdot,\alpha)$ is that

(2.15) $$\alpha = G_L(p) := p^2 f_L'(p) / f_L(p);$$

similarly, $C_U(\cdot,\alpha)$ takes its minimum at p only if

(2.16) $\qquad \alpha = G_U(p) := p^2 f_U'(p) / f_U(p)$.

Sufficient conditions for the existence and uniqueness of solutions to (2.15) and (2.16) (i.e., for well-defined functional inverses of G_L and G_U) which actually minimize $C_L(\cdot,\alpha)$ and $C_U(\cdot,\alpha)$ are given in

Lemma 2.1: Let f_L and f_U be as above, and suppose that

(2.17) $\quad G_L'(p) > 0$ if $G_L(p) > 0$ and $G_U'(p) > 0$ if $G_U(p) > 0$.

Then G_L and G_U have respective functional inverses $p_L^*, p_U^* : \mathbb{R}^{++} \to \mathbb{R}^{++}$ such that for all $p \in \mathbb{R}^{++}$

(2.18) $\qquad C_L^*(\alpha) := C_L(p_L^*(\alpha),\alpha) \le C_L(p,\alpha)$

and

(2.19) $\qquad C_U^*(\alpha) := C_U(p_U^*(\alpha),\alpha) \le C_U(p,\alpha)$

with equality in (2.18) or (2.19) if and only if $p = p_L^*(\alpha)$ or $p = p_U^*(\alpha)$, respectively.

Proof: See Theorem 2.1 and Lemma 3.1 of Werschulz [76]. ∎

We call $p_L^*(\alpha)$ (respectively, $p_U^*(\alpha)$) the lower (upper) optimal order, $C_L^*(\alpha)$ (respectively, $C_U^*(\alpha)$) the lower (upper) optimal complexity, and

(2.20) $h_L^*(\alpha) := h_L(p_L^*(\alpha),\alpha)$ (respectively, $h_U^*(\alpha) := h_U(p_U^*(\alpha),\alpha)$)

the lower (upper) optimal step-size. Combining (2.11), (2.12), and Lemma 2.1, we have

Theorem 2.2:

$$C_L^*(\alpha) \le C^*(\alpha) \le C_U^*(\alpha). \quad ∎$$

We next describe the behavior of these quantities as α increases and tends to infinity.

Theorem 2.3: Let f_L and f_U be as in Lemma 2.1. Then $p_L^*(\alpha)$, $p_U^*(\alpha)$, $C_L^*(\alpha)$, and $C_U^*(\alpha)$ all increase monotonically and tend to infinity with α.

Proof: See Theorems 2.2 and 3.3 of Werschulz [76]. ∎

Finally, we need a restriction of the problem class $\mathcal{D}\times\mathcal{V}$ to "sufficiently difficult" problems; this will allow us to determine $\varkappa_L$ and thus establish lower bounds. We will assume that

$$\sigma_G(\varphi_p,h) \geq (M_L\, h)^p \quad \text{if } h \in I \text{ and } p \in \mathbb{Z}^{++} \tag{2.23}$$

for some $M_L > 0$ independent of h and p. In the methods we study, (2.23) holds provided all sharp upper bounds are attained.

3. Taylor Series Methods

The class Φ_T of Taylor series methods is defined by expanding x in a truncated Taylor series. Thus the increment function φ_p is given by

$$\varphi_p(x_i,h) := \Sigma_{k=0}^{p-1} v^{(k)}(x_i)\, h^k / (k+1)! , \tag{3.1}$$

where

$$v^{(k)}(x_i) := (d/dt)^k [v(x(t))] \Big|_{x(t) = x_i} . \tag{3.2}$$

The usual method of computing (3.2), as described in "classical" numerical analysis texts such as Henrici [62], invokes the chain rule. This quickly leads to expressions of horrifying complexity; for this reason, most texts quickly abandon the discussion of high-order Taylor series methods.

We are interested in faster algorithms for computing φ_p. First, we address the problem of a lower bound for the combinatory cost d(p).

Proposition 3.1: There exists a constant $a_L > 0$ such that any sequence of algorithms for computing Φ_T must satisfy

$$d(p) \geq a_L\, p^N . \tag{3.3}$$

Proof: Any algorithm for computing φ_p requires the information

$$\mathfrak{N}_p(v) := \{D^\beta v:\ 0 \leq |\beta| \leq p - 1\} .$$

(We use the standard multi-index notation found in Friedman [69].) It is then easy to see that the above set has $O(p^N)$ (as $p \uparrow \infty$) distinct elements, which are (generally) independent; this is an immediate consequence of Problem 11 in Chapter 1 of Pólya and Szegö [25]. Thus (3.3) gives a linear lower bound. ∎

Note that the constant a_L in (3.3) depends on N. Since we are treating the case where N is fixed and p is allowed to vary, we will not indicate this dependence explicitly. We now see how close we can get to an optimum value for d(p).

Theorem 3.1: There exists a constant $a_U > 0$ such that the combinatory cost d(p) of computing $\varphi_p \in \Phi_T$ satisfies the bound

$$d(p) \le a_U p^N \ln(p+e) . \tag{3.4}$$

Proof: We first consider the case N = 1. Note that x(h) is the zero of

$$F(z) := \int_{x_0}^{z} d\xi / v(\xi) - h . \tag{3.5}$$

As in Brent and Kung [76], we consider the formal power series

$$P(s) := F(x_0+s) - F(x_0) ,$$

where s is an indeterminate. Let V be the power series reversion of P. Adopting the notation of Brent and Kung [76], we see that

$$x(s) = x_0 + V(s) = x_0 + V_p(s) + O(s^{p+1}) .$$

By the uniqueness of the Taylor coefficients of an analytic function, we see that

$$\varphi_p(x_0,h) = h^{-1} V_p(h) .$$

Since the number $V_p(h)$ can be computed in O(p ln p) operations from the Taylor coefficients of v (by Theorem 6.2 of Brent and Kung [76]), the result for N = 1 follows.

For $N \ge 2$, we use Newton's method (Rall [69]) applied to the formal power series operator P given by

$$(Py)(s) := y(s) - x_0 - \int_0^s v(y(\tau))\, d\tau ;$$

clearly, the formal power series x(s) is the zero of P. The algorithm itself is defined recursively. Let a formal power series $x_{(p)}(s)$ satisfying

$$x_{(p)}(s) = x(s) + O(s^{p+1})$$

be given. Precompute

$$w(s) := \int_0^s v(x_{(p)}(\tau))\, d\tau - x_0 - x_{(p)}(s) + O(s^{2p+2}) , \tag{3.6}$$

$$Q(s) := \nabla v(x_{(p)}(s)) + O(s^{2p+2}) , \tag{3.7}$$

and let $u_{(0)}(s) := 0$. Then set

$$x_{(2p+1)}(s) := x_{(p)}(s) + u_{(p+1)}(s) ,$$

where

$$(3.8)\quad u_{(k+1)}(s) := \int_0^s Q(\tau)\, u_{(k)}(\tau)\, d\tau + w(s) + O(s^{2p+2}),\quad 0 \le k \le p .$$

Following the proof given in Rall [69], we find that

$$x_{(2p+1)}(s) = x(s) + O(s^{2p+2}) .$$

We need only consider the cost T(p,N) of computing the series $x_{(p)}(s)$ in determining d(p), since x(h) may be recovered from the formal power series in O(p) operations. Clearly, we have the recursion

$$(3.9)\qquad T(2p+1,N) \le T(p,N) + T_6 + T_7 + T_8 ,$$

where T_m is the cost of step (3.m) for m = 6, 7, 8. Let COMP(p,N) be the time required to find the first p terms of the formal power series $f(y_1(s), \ldots, y_N(s))$, where $f, y_1, \ldots, y_N$ are formal power series, and $y_1, \ldots, y_N$ have zero constant term. Theorem 7.1 of Brent and Kung [76] states that

$$COMP(p,2) = O(p^2 \ln p) ,$$

and it is easy to show that for any $N \in \mathbb{Z}^{++}$,

$$COMP(p,N+1) = O(p\ COMP(p,N)) .$$

Thus for $N \ge 2$, we have

$$(3.10)\qquad COMP(p,N) = O(p^N \ln p) ,$$

and so we see that

$$T_6 + T_7 = O((2p+1)^N \ln p) .$$

Finally, let MULT(p) be as in Brent and Kung [76]; we see that

$$T_8 = (p+1)\,[N^2\, MULT(2p+1) + O(p)] = O((2p+1)^2 \ln p)$$

if Fast Fourier Transform multiplication (Borodin and Munro [75]) is used. Since $N \ge 2$, we have

$$(3.11)\qquad T_6 + T_7 + T_8 = O((2p+1)^N \ln p) ,$$

and so (3.9) and (3.11) imply that

$$T(p,N) = O(p^N \ln p),$$

which completes the proof. ∎

(Note that the second algorithm is inferior to the first algorithm when applied to the scalar case N = 1, where we find that the second algorithm requires $O(p^2 \ln p)$ arithmetic operations.)

We now determine bounds on $C(p,\alpha)$. First, consider lower bounds. Clearly, there exists $e_L(v) \geq 0$ such that

$$e(D^\beta v_i) \geq e_L(v) \quad (1 \leq i \leq n, |\beta| \in \mathbb{Z}^+) . \tag{3.12}$$

Since $\mathfrak{N}_p(v)$ has $O(p^N)$ elements, there exists a constant $b_L > 0$ such that

$$e(\mathfrak{N}_p(v)) \geq b_L e_L(v) p^N . \tag{3.13}$$

From (3.3) and (3.13), we have a lower-bound cost per step of

$$c_L(p) = [a_L + b_L e_L(v)] p^N . \tag{3.14}$$

This leads to

Theorem 3.2: $C_L(p,\alpha) = M_L [a_L + b_L e_L(v)] p^N e^{\alpha/p}$.

Proof: This is an immediate consequence of (2.23) and (3.14). ∎

Note that $f_L(p) := M_L c_L(p)$ satisfies the conditions of Lemma 2.1. Thus, the optimality theory of Section 2 holds. In particular, we have

Theorem 3.3: $C_L^*(\alpha) = M_L [a_L + b_L e_L(v)] (e/N)^N \alpha^N$.

Proof: From (2.18) and (3.14), we find that $G_L(p) = Np$, so that

$$p_L^*(\alpha) = \alpha/N \quad \text{and} \quad h_L^*(\alpha) = (M_L e^N)^{-1} .$$

The result follows by letting $p = p_L^*(\alpha)$ in the definition of $C_L(p,\alpha)$. ∎

However, recall that we assumed that the non-identical mixed partial derivatives of v are independent. There are a number of systems for which this is not true (for instance, constant coefficient linear systems); for such systems, it is clear that we may

be able to use the extra information of non-independence to find algorithms that are faster than the lower bounds given above. However, we will ignore this case and only consider the problem for a "general" function v.

Next, we turn to upper bounds on the complexity. Theorem 3.1 tells us how to combine the necessary information to get the solution at a new grid-point; we need only measure the cost of getting the information. So, let

$$e^{(k)}(v) = \max \{e(D^{\beta} v_i): \ 1 \le i \le N, \ |\beta| = k\} .$$

Using the result in Pólya and Szegö [25], we see that

$$e(\mathfrak{N}_p(v)) \le N \Sigma_{k=0}^{p-1} e^{(k)}(v) \, (N+k-1)! \, / \, [k!(N-1)!] . \tag{3.15}$$

Unfortunately, the right-hand side of (3.15) does not fit our general model, so we must assume that we know how $e^{(k)}(v)$ changes as k increases. We will consider the case where the cost of derivative evaluation is bounded; that is, we will assume that

$$e^{(k)}(v) \le e_U(v) \tag{3.16}$$

for some $e_U(v)$ independent of k. Other cases (e.g., $e^{(k)}(v) = O(k^m)$ for some $m > 0$) may be analyzed in a similar manner; of course, they will give different results. By (3.15) and (3.16), there is a $b_U > 0$ such that

$$e(\mathfrak{N}_p(v)) \le b_U \, e_U(v) p^N . \tag{3.17}$$

From (3.4) and (3.17), we have an upper-bound cost per step of

$$C_U(p) = a_U \, p^N \ln (p+e) + b_U \, e_U(v) p^N . \tag{3.18}$$

This leads to

Theorem 3.4: There exists an $M_U > 0$ such that

$$C_U(p,\alpha) = M_U [a_U \, p^N \ln (p+e) + b_U \, e_U(v) p^N] \, e^{\alpha/p} .$$

Proof: By Cauchy's Integral Theorem (Ahlfors [66], pg. 122), there exists a $B > 0$ such that

$$|||x^{(k+1)}||| \, / \, (k+1)! \le B^k ,$$

where we define

(3.19) $$|||y||| := \max_{t \in I} \|y(t)\|$$

for any $y: I \to \mathbb{R}^N$. Thus by Section 3.3-3 of Henrici [62], we see that a Lipschitz constant for φ_p in Φ_T is given by

$$\Sigma_{k=0}^{p-1} |||x^{(k+1)}||| \, h^k / (k+1)! \;\le\; \Sigma_{k=0}^{p-1} (Bh)^k \;\le\; L := (1 - Bh_0)^{-1},$$

provided that $h \le h_0 < B^{-1}$. By Section 3.3-2 and 3.3-4 of Henrici [62], there exists an $M_U > 0$ such that

$$\sigma_G(\varphi_p, h) \le (M_U h)^p .$$

The result now follows from Theorem 4.1 and (3.18). ∎

We are now ready to consider the optimal p for $C_U(p,\alpha)$.

Theorem 3.5:

(1.) For all $\alpha > 0$, there exists $p_U^*(\alpha)$ such that (2.19) holds.

(2.) $p_U^*(\alpha)$ increases monotonically with α, and

$$p_U^*(\alpha) \sim \alpha/N \quad \text{as } \alpha \uparrow \infty .$$

(3.) $C_U^*(\alpha)$ increases monotonically with α, and

$$C_U^*(\alpha) \sim M_U a_U (e/N)^N \alpha^N \ln \alpha \quad \text{as } \alpha \uparrow \infty .$$

(4.) $h_U^*(\alpha) \sim (M_U e^N)^{-1}$ as $\alpha \uparrow \infty$.

Proof: Clearly c_U satisfies (2.13) and (2.14). Now write

$$G_U(p) = G_1(p) + G_2(p),$$

where

$$G_1(p) = N p \quad \text{and} \quad G_2(p) = \nu p^2 / D_2(p);$$

here we set

$$D_2(p) := (p+e)\,[(p+e)\ln(p+e) + 1] \quad \text{and} \quad \nu := a_U / [b_U e_U(\nu)] .$$

We see immediately that G_1 satisfies (2.17); a straightforward calculation shows that

$$G_2'(p) = \nu\,[D(p)]^{-2} \{\nu p\,[\ln(p+e)] - 1] + 2e[\nu \ln(p+e) + 1]\},$$

so that $G_2'(p) > 0$ for $p > 0$. Thus G_2 satisfies (2.17), which shows that G_U satisfies (2.17). Hence p_U^* and C_U^* behave as described in Theorem 2.2.

Since $p_U^*(\alpha)$ goes to infinity with α, we see that

$$\alpha = G_U(p_U^*(\alpha)) \sim N\, p_U^*(\alpha) + p_U^*(\alpha) / \ln p_U^*(\alpha) \sim N\, p_U^*(\alpha),$$

which gives the asymptotic estimate in (2.). The rest of the Theorem follows from this estimate. ∎

Unfortunately, the estimates given above are only asymptotic as $\alpha \uparrow \infty$; this will be typical, since many of the equations to be solved involve products of logarithmic and polynomial terms, and thus cannot be solved exactly. On the other hand, these asymptotic expressions are sufficient for our purposes, since they describe how quickly $p_U^*(\alpha)$ and $C_U^*(\alpha)$ increase with α.

Note that as α tends to infinity, $C_U^*(\alpha)$ becomes independent of $e_U(v)$, which measures how hard it is to evaluate the derivatives of v; this is because the combinatory cost eventually overwhelms the informational cost. This kind of behavior will be typical of the complexity analyses in this paper. Finally, note that the bound

$$(3.20)\quad C_L^*(\alpha) = O(\alpha^N) \le C^*(\alpha) \le O(\alpha^N \ln \alpha) = C_U^*(\alpha) \text{ as } \alpha \uparrow \infty$$

implies that

$$C_U^*(\alpha) / C_L^*(\alpha) = O(\ln \alpha) \text{ as } \alpha \uparrow \infty;$$

this indicates the gap in our knowledge of the complexity of solving (1.1) via Taylor series methods.

4. Linear Runge-Kutta Methods

For many functions v, caculation of the derivatives required by Taylor series methods is prohibitively expensive. For this reason, we are interested in methods which use information that is somewhat more readily available. In particular, we will consider methods that use only evaluations of v, combined in a highly structured manner. We say that Φ_{LRK} is a class of linear Runge-Kutta methods (abbreviated, "LRK methods") if each increment function φ_p may be written in the form

$$\varphi_p(x_i,h) := \Sigma_{l=0}^{s-1} \lambda_{sl} k_l \tag{4.1}$$

where

$$k_l := v(x_i + h \Sigma_{j=0}^{l-1} \lambda_{lj} k_j) \quad \text{for } 0 \le l \le s-1 , \tag{4.2}$$

the integer $s = s(p)$ is said to be the number of stages of φ_p; the number of stages is equal to the number of times the vector function v must be evaluated. (In order to simplify notation, we will not explicitly indicate the dependence of λ_{lj} and k_j on p.) The method φ_p defined by (4.1) and (4.2) is explicit in that k_l depends only on $k_0, \dots, k_{l-1}$; see Butcher [64] for a discussion of semi-explicit and implicit methods. (We use the adjective "linear" to distinguish these methods from "nonlinear Runge-Kutta methods," which were first proposed in Brent [74].)

Since the function φ_p is (in practice) always evaluated by using the obvious algorithm suggested by its definition, we shall identify an algorithm for evaluating φ_p with φ_p itself. Thus the problem of finding the best algorithm for evaluating φ_p in Φ_{LRK} is equivalent to the problem of finding the best basic sequence of LRK methods possible. This is related to the problem of finding the smallest value of s(p) such that φ_p has order p. This minimal value is given by

$$(4.3) \qquad s(p) = \begin{cases} p & p = 1, 2, 3, 4 \\ p+1 & p = 5, 6 \\ p+2 & p = 7 \\ \text{unknown} & p \geq 8 \end{cases}$$

For methods of order greater than seven, a gap develops. For instance, eighth-order methods with eleven stages exist, and it is known that any eighth-order method requires at least ten stages. For arbitrary $p \geq 8$, the best bounds known for the optimum value of s(p) are

$$(4.4) \qquad p + \theta(p) \leq s(p) \leq (p^2 - 7p + 14) / 2 ,$$

where $\theta(p) \geq c \ln p$ for all sufficiently large p (for some $c > 0$). The lower bound is given in Butcher [75]; the proof is quite involved, and the result is not much better than the "trivial" lower bound $s(p) \geq p$ (Hindmarsh [74], page 84). A class Φ_{CVRK} of methods such that φ_p requires only $(p^2 - 7p + 14) / 2$ stages is given in Cooper and Verner [72].

We first consider lower bounds on the complexity $C(p,\alpha)$ using LRK methods. The "trivial" lower bound $s(p) \geq p$ will be used, since the term $\theta(p)$ will be small when p is small and will not affect the asymptotic behavior of optimal order and complexity for p large. It is known (Butcher [64]) that at least $O(p^2)$ of the subdiagonal elements of the matrix A (whose elements are the λ_{ij} in (4.2)) must be non-zero in order for A to define a p^{th}-order method. Thus there exists $a_L > 0$ such that

$$(4.5) \qquad d(p) \geq a_L p^2 ;$$

since $s(p) \geq p$, we see that

$$(4.6) \qquad e(\mathfrak{R}_p(v)) \geq N e_L(v) p ,$$

where we now write

$$e_L(v) := \min_{1 \leq i \leq N} e(v_i) .$$

Thus (4.5) and (4.6) show that a lower bound on the cost per step for φ_p is given by

$$c_L(p) = a_L p^2 + N e_L(v)\, p \ . \tag{4.7}$$

Theorem 4.1:

$$C_L(p,\alpha) = M_L [a_L p^2 + N e_L(v)\, p]\, e^{\alpha/p} \ .$$

Proof: This follows immediately from (2.23) and (4.7). ∎

It is clear that $f_L(p) := M_L [a_L p^2 + N e_L(v)\, p]\, e^{\alpha/p}$ satisfies (2.13) and (2.14). We claim that f_L yields a G_L satisfying (2.17). Indeed, write

$$f_L(p) = f_1(p)\, f_2(p) \ ,$$

where

$$f_1(p) := M_L a_L p$$

and

$$f_2(p) := p + \nu, \quad \text{where} \quad \nu := N e_L(v) / a_L \ .$$

Clearly f_1 yields a G_1 satisfying (2.17). Since f_2 is a linear polynomial with a negative zero, it may be shown that f_2 yields a G_2 satisfying (2.17). Thus f_L yields a G_L satisfying (2.17); in fact, we have

$$G_L(p) = G_1(p) + G_2(p) = p\,[1 + (1 + \nu p^{-1})^{-1}] \ . \tag{4.8}$$

This leads us to

Theorem 4.2:

$$C_L^*(\alpha) \sim [M_L a_L e^2 / 4]\, \alpha^2 \quad \text{as} \quad \alpha \uparrow \infty \ .$$

Proof: From (4.8), we see that $G_L(p) \sim 2p$ as $p \uparrow \infty$. Since (2.13), (2.14), and (2.17) hold, $p_L^*(\alpha)$ tends to infinity with α. Thus

$$\alpha = G_L(p_L^*(\alpha)) \sim 2\, p_L^*(\alpha) \quad \text{as} \quad \alpha \uparrow \infty \ ,$$

i.e., $p_L^*(\alpha) \sim \alpha/2$ as $\alpha \uparrow \infty$. The result now follows from Theorem 4.1. ∎

We now turn to upper bounds on complexity. The class Φ_{CVRK} derived in

Cooper and Verner [72] has two deficiencies, the first of which is that no uniform upper bound on $\sigma_{LU}(\varphi_p,h)$ is known for Φ_{CVRK}; in addition, the combinatory cost for this class of methods is $O(p^4)$ as $p \uparrow \infty$. Instead, we turn to the basic sequence Φ_{CRK} discussed in the Appendix. There, we prove that there is an $M_U > 0$ such that

$$\sigma_G(\varphi_p,h) \leq (M_U \ln (p + e)\, h)^p , \tag{4.9}$$

provided $h \leq h_p$, where $h_p = O((\ln p)^{-1})$ as $p \uparrow \infty$. Furthermore, there are a large number of extra zeros in the matrix A for $\varphi_p \in \Phi_{CRK}$. Using the notation of the Appendix, we see that the number of non-zero entries in A is

$$\begin{aligned} \Sigma_{i=0}^{s} \xi_i &= \Sigma_{i=1}^{p-1} i^2 + p \\ &= p^3/3 - p^2/2 + 7p/6 \\ &\leq p^3/3 + 2p^2/3 \end{aligned}$$

for $p \in \mathbb{Z}^{++}$. Finally, note that the number of stages $s(p)$ required for $\varphi_p \in \Phi_{CRK}$ is

$$s(p) = \lfloor (p^2 - 2p + 4)/2 \rfloor \leq p^2/2 + p \tag{4.10}$$

for $p \in \mathbb{Z}^{++}$, which shows that the number of stages required for a p^{th}-order method in Φ_{CRK} asymptotically equals the number requires for a p^{th}-order method in Φ_{CVRK}. Thus (considering the combinatory costs), the class Φ_{CVRK} actually costs more per step than does Φ_{CRK}; ignoring the combinatory costs would have caused us to reach the opposite conclusion.

First, we look at the cost per step. By (4.10),we see that

$$e(\mathfrak{R}_p(v)) \leq \frac{1}{2} (p^2 + p)\, N\, e_U(v) , \tag{4.11}$$

where

$$e_U(v) := \max_{1 \leq i \leq N} e(v_i) .$$

Since we are using Φ_{CRK}, it is easy to see that there is a $b_U \geq 2/3$ such that

$$d(p) \leq (p^3/3 + b_U p^2) \cdot 2N . \tag{4.12}$$

Combining (4.11) and (4.12), we see that the total combinatory cost per step is bounded by

$$c_U(p) = N [\, 2p^3/3 + \beta_1 p^2 + \beta_2 p \,], \tag{4.13}$$

where

$$\beta_1 := e_U(v) / 2 + 2 b_U \quad \text{and} \quad \beta_2 := e_U(v) / 2 .$$

Using (4.9) and (4.13) gives

Theorem 4.3:

$$C_U(p,\alpha) = M_U N [2p^3/3 + \beta_1 p^2 + \beta_2 p \,] \ln (p + e) e^{\alpha/p} . \blacksquare$$

Now we look at the optimality theory for the upper bound.

Theorem 4.4.:

(1.) For all $\alpha > 0$, there exists $p_U^*(\alpha)$ such that (2.19) holds.

(2.) $p_U^*(\alpha)$ increases monotonically with α, and

$$p_U^*(\alpha) \sim \alpha/3 \quad \text{as } \alpha \uparrow \infty .$$

(3.) $C_U^*(\alpha)$ increases monotonically with α, and

$$C_U^*(\alpha) \sim [\, 2 M_U N e^3 / 81 \,] \alpha^3 \ln \alpha \quad \text{as } \alpha \uparrow \infty.$$

(4.) $h_U^*(\alpha) \sim (M_U e^3 \ln \alpha)^{-1}$ as $\alpha \uparrow \infty$.

Proof: We write

$$f_U(p) := M_U \ln (p + e) c_U(p)$$

in the form

$$f_U(p) = f_1(p) f_2(p) ,$$

where

$$f_1(p) = M_U N p \ln (p + e) \quad \text{and} \quad f_2(p) = 2p^2/3 + \beta_1 p + \beta_2 .$$

It is clear that f_1 satisfies the hypotheses of Lemma 2.1. Now we consider f_2. Clearly f_2 has no positive zeros; it may be seen that the condition $b_U \geq 2/3$ implies that f_2 has

a positive discriminant and hence has no complex roots. Thus f_2 has only negative roots; one may then show that this guarantees that f_2 satisfies the hypotheses of Lemma 2.1. Thus, the same may be said for $f = f_1 f_2$.

Thus p_U^* and C_U^* behave as described in (1.) of Theorem 2.3. We also see that $G_U(p) \sim 3p$ as $p \uparrow \infty$. Thus the estimate in (2.) holds, from which we get the estimates in (3.) and (4.). ∎

So in the class of linear Runge-Kutta methods, we find that

$$C_L^*(\alpha) = O(\alpha^2) \leq C^*(\alpha) \leq C_U^*(\alpha) = O(\alpha^3 \ln \alpha) \tag{4.14}$$

as α tends to infinity; hence, the ratio

$$C_U^*(\alpha) / C_L^*(\alpha) = O(\alpha \ln \alpha)$$

indicates the gap in our knowledge of the complexity of linear Runge-Kutta methods.

5. Comparison of the Methods

We now wish to compare the classes of Taylor series methods and LRK methods. Write $C_{U,T}^*$, $C_{L,T}^*$, and C_T^* (respectively, $C_{U,LRK}^*$, $C_{L,LRK}^*$, and C_{LRK}^*) for C_U^*, C_L^*, and C^* in the class Φ_T (respectively, the class Φ_{LRK}). Since we have only asymptotic expressions for these quantities, we are forced to use an asymptotic comparison. If $f, g : \mathbb{R}^{++} \to \mathbb{R}^{++}$ satisfy $\lim_{\alpha \uparrow \infty} f(\alpha) = \lim_{\alpha \uparrow \infty} g(\alpha) = +\infty$, we will write

$$f \prec g \quad \text{iff} \quad f(\alpha) = o(g(\alpha)) \text{ as } \alpha \uparrow \infty ; \tag{5.1}$$

we say that f is asymptotically less than g. If $f \prec g$, there is an $\alpha_0 > 0$ such that $f(\alpha) < g(\alpha)$ for $\alpha > \alpha_0$, so there is a non-asymptotic interpretation of the order relation $\prec$. Thus if f and g are cost functions, the statement "$f \prec g$" implies that the method whose cost is given by f is "better" (i.e., cheaper) than the method whose cost is given by g, for ε sufficiently small. Using the results of (3.20) and (4.14), we then have the following

Theorem 5.1: Suppose that (3.16) holds.

(1.) If $N = 1$, then $C_{U,T}^* \prec C_{L,LRK}^*$.

(2.) If $N = 2$, then $C_{U,T}^* \prec C_{U,LRK}^*$.

(3.) If $N = 3$, then

$$C_{U,T}^*(\alpha) = O(C_{U,LRK}^*(\alpha))$$

and

$$C_{U,LRK}^*(\alpha) = O(C_{U,T}^*(\alpha))$$

as $\alpha \uparrow \infty$.

(4.) If $N \geq 4$, then $C_{U,LRK}^* \prec C_{L,T}^*$. ∎

If (3.16) does not hold, then (1.), (2.), and (3.) may be false, but (4.) will certainly be true. As an immediate corollary to the above theorem, we have

Theorem 5.2:

(1.) If $N = 1$ and (3.16) holds, then $C_T^* < C_{LRK}^*$.

(2.) If $N \geq 4$, then $C_{LRK}^* < C_T^*$. ∎

So if the derivatives of v are cheap to evaluate, we see that the best Taylor series method known is better than the best linear Runge-Kutta method possible for the scalar case $N = 1$; but if $N \geq 4$, the best linear Runge-Kutta method known is better than the best Taylor series method possible.

Appendix: Error Bounds for a Sequence of LRK Methods

In this Appendix, we describe a subclass of a class of linear Runge-Kutta ("LRK") methods due to Cooper [69]. We shall first prove the following

Theorem A.1: There is a basic sequence $\Phi_{CRK'}$ of LRK methods such that

(1.) Each $\varphi_p \in \Phi_{CRK'}$ requires

$$s(p) := (p^2 - p + 2) / 2$$

evaluations of v per step.

(2.) There exists an $M_U > 0$ such that

$$\sigma_G(\varphi_p, h) \leq (M_U \ln (p+e)\, h)^p \tag{A.1}$$

for $h \leq h_p = O((\ln p)^{-1})$.

We use the notation of Cooper and Verner [72]. Let $p \in \mathbb{Z}^{++}$ be given; define $\rho: \mathbb{Z}^+ \cap [0, p] \to \mathbb{Z}^+$ by

$$\rho(j) := \begin{cases} \Sigma_{k=0}^{j} k = j(j+1) / 2 & \text{if } j \neq p \\ s & \text{if } j = p \ , \end{cases} \tag{A.2}$$

where we write "s" for "s(p)" as defined above. Next, a set $\{\xi_0, \ldots, \xi_s\}$ of integers is defined by picking $\xi_0 := p$, and setting ξ_i $(i \neq 0)$ to be the unique integer in $[1, p]$ satisfying

$$\rho(\xi_i - 1) < i \leq \rho(\xi_i) \ . \tag{A.3}$$

We now pick $u_0, \ldots, u_s \in I$ satisfying

$$u_0 = 0 \ , \ u_s = 1 \ , \ u_i \neq 0 \text{ if } i \neq 0 \tag{A.4}$$

and

$$(\xi_i = \xi_j \text{ and } i \neq j) \text{ implies } u_i \neq u_j \ . \tag{A.5}$$

Finally, we pick a matrix of coefficients $A := \{\lambda_{ij}: 0 \leq j \leq i-1, 1 \leq i \leq s\}$ such that

$$\lambda_{ij} = 0 \quad \text{if} \quad \xi_i < \xi_j - 1 \qquad (1 \leq i, j \leq s) \tag{A.6}$$

and

$$(A.7)\qquad \Sigma_{j=0}^{i-1} \lambda_{ij}\, u_j^{\tau} = (\tau+1)^{-1} u_i^{\tau+1} \qquad (0 \le \tau \le \xi_i - 1,\ 1 \le i \le s)\ .$$

Cooper and Verner [72] point out that these conditions may always be fulfilled; the resulting Λ defines a p^{th}-order LRK method with s stages.

We are interested in a choice of $u_0, \dots, u_s$ which will give a small error coefficient. To this end, we will choose

$$(A.8)\qquad \{u_j:\ \xi_j = n\} = \{(1 + x_{kn}) / 2 :\ 1 \le k \le n\} \qquad (1 \le n \le p - 1)\,,$$

where $x_{1n}, \dots, x_{rn}$ are the zeros of the Jacobi polynomial $P_n := P_n^{(1,1)}$ (see Szegő [59]). Since these zeros are distinct and lie in $[-1, 1]$, conditions (A.4) and (A.5) may be satisfied.

Now we are able to exhibit a solution to the i^{th} system in (A.7). First, note that the equation for $\tau = 0$ may be separated from the others, since $u_0 = 0$. Setting

$$n := \xi_i - 1\,,$$

we see that

$$(A.9)\qquad \lambda_{i0} = u_i - \Sigma_{j=1}^{i-1} \lambda_{ij} = u_i - \Sigma \{ \lambda_{ij}:\ j < i \text{ and } \xi_j \ge n \}\,,$$

the last by (A.6). We wish to determine the nonzero λ_{ij}, i.e., those λ_{ij} for which $\xi_j \ge n$ and $j < i$. So setting

$$\lambda_{ij} = 0 \quad \text{unless} \quad j \in \{j_1, \dots, j_n\}\,,$$

we see that the remaining λ_{ij} are the solution of the system

$$\Sigma_{k=1}^{n} u_{j_k}^{\tau} \lambda_{ij_k} = (\tau+1)^{-1} u_i^{\tau+1} \qquad (1 \le \tau \le n)\ .$$

Thus the λ_{ij_k} are the weights for an interpolatory quadrature formula on $[0, u_i]$ with abscissae $u_{j_1}, \dots, u_{j_n}$. From the usual expression for such weights and (A.6), we see that

$$\lambda_{ij_k} = \mu_{ikn} := [2P_n'(\cos \vartheta_{kn})]^{-1} \int_{\vartheta_{i,n+1}}^{\pi} [P_n(\cos \vartheta) / (\cos \vartheta - \cos \vartheta_{kn})] \sin \vartheta \, d\vartheta\,,$$

where $x_{kn} = \cos \vartheta_{kn}$ $(1 \le k \le n)$.

Lemma A.1: $\mu_{ikn} = O(n^{-1} \ln n)$ as $n \uparrow \infty$.

Proof: Since the zeros of P_n are symmetric about the origin, we may assume that $0 < \vartheta_{kn} \le \pi/2$. Using (8.9.2) of Szegő [59], we then find

$$\mu_{ikn} = O(k^{5/2} n^{-3}) \int_{\vartheta_{i,n+1}}^{\pi} [P_n(\cos\vartheta) / (\cos\vartheta - \cos\vartheta_{kn})] \sin\vartheta \, d\vartheta \ .$$

Case 1: $\vartheta_{1,n+1} \le \vartheta_{i,n+1} \le \vartheta_{k,n+1}/2$. We consider the integral over $[\vartheta_{1n}/2, \vartheta_{i,n+1}]$, since Theorem 15.4 of Szegő [59] proves that

$$O(k^{5/2} n^{-3}) \left[\left| \int_0^{\pi} \right| + \left| \int_0^{\vartheta_{1n}/2} \right| \right] = O(n^{-1}) \ .$$

(Here the integrand is the same as in the preceding integral.) But the proof of (15.4.12) in Szegő [59] extends almost immediately to a proof that the remaining integral is $O(k^{-2} n)$, since (15.4.12) is proved by order-of-magnitude estimates. Thus $\mu_{ikn} = O(n^{-1}) = O(n^{-1} \ln n)$ for Case 1.

Case 2: $\vartheta_{k,n+1}/2 \le \vartheta_{i,n+1} \le 3\vartheta_{k,n+1}/2$. We consider the integral over $[\vartheta_{kn}/2, \vartheta_{i,n+1}]$, since Szegő [59] shows that

$$O(k^{5/2} n^{-3}) \left| \int_{\vartheta_{kn}/2}^{\pi} \right| = O(n^{-1}) \ .$$

As in (15.4.13) of Szegő [59], we have

$$\int_{\vartheta_{kn}/2}^{\vartheta_{i,n+1}} = O(nk^{-3/2}) \, I_1 + I_2 \ .$$

Here

$$I_1 := \int_{\vartheta_{kn}/2}^{\vartheta_{i,n+1}} D(\vartheta) \sin\vartheta \, d\vartheta \ ,$$

with

$$D(\vartheta) := [\cos(N\vartheta + \gamma) - \cos(N\vartheta_{kn} + \gamma)] / [\cos\vartheta - \cos\vartheta_{kn}] \ ,$$

where $N := n + 3/2$ and $\gamma := -3\pi/4$, and

$$I_2 := \int_{\vartheta_{kn}/2}^{\vartheta_{i,n+1}} R_n(\vartheta, \vartheta_{kn}) \sin\vartheta \, d\vartheta = O(nk^{-3/2}) \ ,$$

with R_n the remainder term in (8.8.2) of Szegő [59]. Unfortunately, the proof that (15.4.14) of Szegő [59] is bounded does not extend to a proof that I_1 is bounded,

since the proof of the former requires that the interval of integration be symmetric about ϑ_{kn}. However, it is straightforward to verify that

$$I_1 = O(1) \int_0^{\pi/4} |\sin N\vartheta / \vartheta| \, d\vartheta = 0 (\ln n) .$$

Thus $\mu_{ikn} = O(n^{-2}k \ln n) = O(n^{-1} \ln n)$ for Case 2.

Case 3: $3\vartheta_{k,n+1} \le \vartheta_{i,n+1} \le 3\pi/4$. We consider the integral over $[3\vartheta_{kn}/2, \vartheta_{i,n+1}]$, since Szegő [59] proves that

$$O(k^{5/2}n^{-3}) \, | \int_{3\vartheta_{kn}/2}^{\pi} | = O(n^{-1}) .$$

But the proof of (15.4.19) in Szegő [59] extends to prove that the remaining integral is $O(k^{-5/2}n)$ (as in Case 1). Thus $\mu_{ikn} = O(n^{-1}) = O(n^{-1} \ln n)$ for Case 3.

Case 4: $3\pi/4 \le \vartheta_{i,n+1} \le \vartheta_{n+1,n+1}$. We consider the integral over $[3\pi/4, \vartheta_{i,n+1}]$, since Szegő [59] shows that

$$O(k^{5/2}n^{-3}) \, | \int_{3\pi/4}^{\pi} | = O(n^{-1}) .$$

As in Cases 1 and 3, the proof of the above may be extended to prove a similar bound on the integral of interest. Thus $\mu_{ikn} = O(n^{-1}) = O(n^{-1} \ln n)$ in Case 4, completing the proof of the Lemma. ∎

Thus (A.9) and Lemma A.1 show the existence of a $\lambda > 0$ such that

$$\Sigma_{j=0}^{i-1} |\lambda_{ij}| \le \lambda \ln (\xi_i + e) ; \tag{A.10}$$

here λ is independent of p. Moreover, the result for the case i = s may be sharpened. We see that $\lambda_{sj} \ge 0$, since the u_j for the s^{th} system in (A.7) are the abscissae for Lobatto quadrature. Thus

$$\Sigma_{j=0}^{s-1} |\lambda_{sj}| = \Sigma_{j=0}^{s-1} \lambda_{sj} = 1, \tag{A.11}$$

the consistency condition in the last equality being a consequence of (A.7) with $\tau = 0$.

Proof of Theorem A.1: As in Cooper and Verner [72], we define

$$\varepsilon_i := \dot{x}(u_i h) - k_i$$

and

$$\delta_i := \int_0^{u_i} \dot{x}(uh)\,du - \Sigma_{j=0}^{i-1} \lambda_{ij}\, \dot{x}(u_j h)$$

for $0 \le i \le s$; note that $\delta_0 = \varepsilon_0 = 0$. Let $z(h)$ be the computed approximation to $x(h)$; then

$$\begin{aligned} h^{-1}\,\|x(h) - z(h)\| &= \| h^{-1}\,[x(h) - x(0)] - \Sigma_{i=0}^{s-1} \lambda_{si}\, k_i \| \\ &\le \|\delta_s\| + \|\Sigma_{i=0}^{s-1} \lambda_{si}\, \varepsilon_i\| \\ &\le \|\delta_s\| + \max_{\xi_i = p-1} \|\varepsilon_i\|, \end{aligned} \tag{A.12}$$

the last by (A.6) and (A.11). By the analyticity of x, there is an $A_1 > 0$ such that

$$\beta_i := h^{-1}\,\| x(u_i h) - \Sigma_{\tau=0}^{\xi_i} (u_i h)^\tau x^{(\tau)}(0) / \tau! \| \le (A_1 h)^{\xi_i}$$

and

$$\gamma_{ij} := \| \dot{x}(u_j h) - \Sigma_{\tau=0}^{\xi_i - 1} (u_j h)^\tau\, \dot{x}^{(\tau)}0) / \tau! \| \le (A_1 h)^{\xi_i},$$

so that the definition of δ_i gives

$$\begin{aligned} \|\delta_i\| &\le \beta_i + \Sigma_{j=0}^{i-1} |\lambda_{ij}|\, \gamma_{ij} \\ &\le (A_1 h)^{\xi_i} + \Sigma_{j=0}^{i-1} |\lambda_{ij}|\, (A_1 h)^{\xi_i} \\ &\le (A_2 h)^{\xi_i} \end{aligned} \tag{A.13}$$

for a suitable $A_2 > 0$. Thus (A.12) becomes

$$h^{-1}\,\|x(h) - z(h)\| \le (A_2 h)^p + \max_{\xi_i = p-1} \|\varepsilon_i\|\,. \tag{A.14}$$

We now use Lemma 1.1 of Cooper and Verner [72] and (A.6) to find that if L is a Lipschitz constant for v, then there exists $A_3 > 0$ such that

$$\begin{aligned} \|\varepsilon_i\| &\le hL\,\|\delta_i\| + hL\,\Sigma_{j=0}^{i-1} |\lambda_{ij}|\, \max_j \|\varepsilon_j\| \\ &\le (A_3 h)^{\xi_i + 1} + (A_3 h)\,\ln(\xi_i + e)\, \max_j \|\varepsilon_j\|, \end{aligned}$$

the last by (A.10) and (A.13); here, the maximum is taken over all $j < i$ such that $\xi_j \ge \xi_i - 1$. A straightforward induction shows that if $(1 + \ln 2)\, A_3 h < 1$, then

$$\|\varepsilon_i\| \le (A_4 \ln(\xi_i + e)\, h)^{\xi_i + 1}$$

for a suitable $A_4 > 0$. Combining this with (A.14), we find

$$h^{-1}\,\|x(h) - z(h)\| \le (A_5 \ln(p+e)\, h)^p, \tag{A.15}$$

the desired bound for the local error for a single unit step.

To extend (A.15) to a global error result, we must look at the Lipschitz constants for the increment functions. Let L be a bound on $\|\nabla v\|$, and write "$\nabla\varphi_p(y,h)$" to indicate gradient with respect to the vector variable y. Now

$$\|\nabla\varphi_p(y,h)\| \le \Sigma_{i=0}^{s-1} |\lambda_{si}| \max_{0\le i\le s-1} \|\nabla k_i(y,h)\|$$
$$= \max_{0\le i\le s-1} \|\nabla k_i(y,h)\| ,$$

where we write "$k_i(y,h)$" to indicate the dependence of k_i upon y and h. By the definition of $k_i(y,h)$, we find

$$\nabla k_i(y,h) = \nabla v(u) \left[1_{N\times N} + h \Sigma_{j=0}^{i-1} \lambda_{ij} \nabla k_j(y,h)\right] ,$$

where $u := y + h \Sigma_{j=0}^{i-1} \lambda_{ij} k_j(y,h)$ and $1_{N\times N}$ is an N×N identity matrix. Taking norms in the above gives the result

$$\xi_i \le L\lambda + hL\lambda \left[\ln(\xi_i+e) \max \{ \xi_j : j < i \text{ and } \xi_j \ge \xi_i - 1 \} \right] ,$$

where $\xi_i := \|\nabla k_i(y,h)\|$. Writing λ_p for the Lipschitz constant for φ_p, it is easy to see that (A.16) and the above inequality imply

$$\lambda_p \le \Sigma_{j=0}^{p-1} (hL\lambda)^j \Pi_{k=1}^{j-2} \ln(p+e-k) ,$$

which is bounded for all p, provided that $h \le h_p < (L\lambda \ln(p+e))^{-1}$. Thus (A.1) follows from this result, (A.15), and Theorem 3.3 of Henrici [62]. ∎

The value for s(p) indicated in Theorem A.1 may be improved somewhat by noting that since we are using a Lobatto quadrature, higher order may be expected with fewer steps. Indeed, if we use the strategy outlined in the comments following Theorem 4 of Cooper and Verner [72], we have

Theorem A.2: There exists a basic sequence Φ_{CRK} of LRK methods such that (A.1) holds and φ_p requires

$$s(p) := \lfloor (p^2 - 2p + 4) / 2 \rfloor$$

evaluations of v per step. ∎

Acknowledgements

I would like to thank Professor R. P. Brent of the Australian National University, Professors H. T. Kung and J. F. Traub of Carnegie-Mellon University, and Professor H. Woźniakowski of the University of Warsaw for their comments and suggestions on the results reported in this paper. In addition, a number of the results in Sections 3 and 4 were obtained by using the MACSYMA system developed by the Mathlab group at Massachusetts Institute of Technology, which is supported by the Defense Advanced Research Projects Agency work order 2095, under Office of Naval Research Contract N00014-75-C-0661.

References

Ahlfors [66]:
Ahlfors, L. V., Complex Analysis, Second Edition. New York: McGraw-Hill, 1966.

Borodin and Munro [75]:
Borodin, A. and I. Munro, The Computational Complexity of Algebraic and Numeric Problems. New York: American Elsevier, 1975.

Brent [74]:
Brent, R. P., "Efficient Methods for Finding Zeros of Functions Whose Derivatives are Easy to Evaluate," Report, Computer Science Department, Carnegie-Mellon University, 1974.

Brent and Kung [76]:
Brent, R. P. and H. T. Kung, "Fast Algorithms for Manipulating Formal Power Series," Report, Computer Science Department, Carnegie-Mellon University, 1976.

Buck [65]:
Buck, R. C., Advanced Calculus, Second Edition. New York: McGraw-Hill, 1965.

Butcher [64]:
Butcher, J. C., "Implicit Runge-Kutta Processes," Math. Comp., Vol. 18, pp. 50-64, January, 1964.

Butcher [75]:
Butcher, J. C., "An Order Bound for Runge-Kutta Methods," SIAM J. Num. Anal., Vol. 12, No. 3, pp. 304-315, June, 1975.

Cooper [69]:
Cooper, G. J., "Error Bounds for Some Single-Step Methods," Conf. on the Numerical Solution of Differential Equations, Lecture Notes in Mathematics 109, pp. 140-147. Berlin: Springer-Verlag, 1969.

Cooper and Verner [72]:
Cooper, G. J. and J. H. Verner, "Some Explicit Runge-Kutta Methods of High Order," SIAM J. Num. Anal., Vol. 9, No. 3, pp. 389-405, September, 1972.

Friedman [69]:
Friedman, A., Partial Differential Equations. New York: Holt, Rinehart, and Winston, 1969.

Gurtin [75]:
Gurtin, M. E., "Thermodynamics and Stability," Arch. Rat. Mech. Anal., Vol. 59, No. 1, pp. 63-96, 1975.

Henrici [62]:
Henrici, P., Discrete Variable Methods in Ordinary Differential Equations. New York: Wiley, 1962.

Hindmarsh [74]:
Hindmarsh, A. C., "Numerical Solution of Ordinary Differential Equations: Lecture Notes." Lawrence Livermore Laborotory Report No. UCID-16588, June, 1974.

Pólya and Szegö [25]:
Pólya, G. and G. Szegö, Aufgaben und Lehrsätze der Analysis, Vol. I. Berlin: Springer-Verlag, 1925.

Rall [69]:
Rall, L. B., Computational Solution of Nonlinear Operator Equations. New York: John Wiley and Sons, Inc., 1969.

Stetter [73]:
Stetter, H. J., Analysis of Discretization Methods for Ordinary Differential Equations. Berlin: Springer-Verlag, 1972.

Szegö [59]:
Szegö, G., Orthogonal Polynomials. Amer. Math. Soc. Colloquium Publications, Vol. XXIII. New York: Amer. Math. Soc., 1959.

Traub and Woźniakowski [76]:
Traub, J. F. and H. Woźniakowski, "Strict Lower and Upper Bounds on Iterative Complexity," in Analytic Computational Complexity, edited by J. F. Traub. New York: Academic Press, 1976.

Werschulz [76]:
Werschulz, A. G., "Optimal Order and Minimal Complexity of One-Step Methods for Initial Value Problems," Report, Computer Science Department, Carnegie-Mellon University, 1976.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) COMPUTATIONAL COMPLEXITY OF ONE-STEP METHODS FOR SYSTEMS OF DIFFERENTIAL EQUATIONS.	5. TYPE OF REPORT & PERIOD COVERED (9) Interim rept. 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Arthur G. Werschulz	8. CONTRACT OR GRANT NUMBER(s) (15) N00014-76-C-0370, NR 044-422 ✓NSF-MCS 75-222-55
9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Computer Science Dept. Pittsburgh, PA 15213	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217	12. REPORT DATE (11) Sep 76 13. NUMBER OF PAGES 35 (12) 36p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

O((p to the N power) ln p) p goes to plus infinity

20. Abstract: The problem is to calculate an approximate solution of an initial value problem for an autonomous system of N ordinary differential equations. Using fast power series techniques, we exhibit an algorithm for the p^{th}-order Taylor series method requiring only $O(p^N \ln p)$ arithmetic operations per step as $p \to +\infty$. (Moreover, we show that any such algorithm requires at least $O(p^N)$ operations per step.) We compute the order which minimizes the complexity bounds for Taylor series and linear Runge-Kutta methods, and show that in all cases, this optimal order increases as the error criterion ε decreases, tending to infinity as ε tends to zero. Finally, we show → next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

UNCLASSIFIED 403 081

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

cont →

that if certain derivatives are easy to evaluate, then Taylor series methods are asymptotically better than linear Runge-Kutta methods for problems of small dimension N.